DOCUMENT RESUME

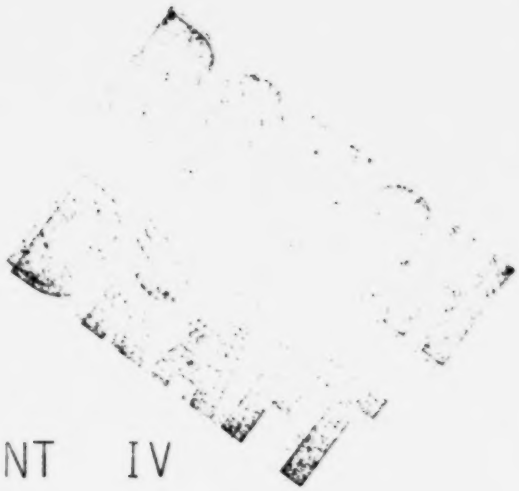
ED 077 973 TM 002 783

TITLE	Procedures for Program Evaluation. Planning Document 4. Rough Draft.
INSTITUTION	Milwaukee Public Schools, Wis.
PUB DATE	Oct 72
NOTE	8p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Criterion Referenced Tests; *Evaluation Methods; *Formative Evaluation; Guides; *Measurement Techniques; *Program Evaluation; Standardized Tests

ABSTRACT

Three types of evaluation used in program evaluation and the type of instrument or measure used are discussed. The three types of evaluation are: comparison group evaluation, analysis of learner outcome or product objectives, and process evaluation. Measurements employed are either direct or indirect. Most measurement in education is usually indirect, by which the subject may be measured while he either actively or passively participates in the measurement act. When it is determined that a test will be employed as a means of indirectly sampling some behavior, two basic kinds of tests are employed--the standardized norm referenced test and a criterion referenced test. Once the evaluation type and instrumentation have been defined, there is the question of the consequences of evaluation. The object of evaluation is to provide decision-making information about programs. Steps in program evaluation are: (1) consideration of how objectives will be evaluated should begin prior to program development; (2) following consultation with evaluation personnel, the planning group should make decisions on what measurement instrument, performance objective formats, and criterion levels represent fair and reasonable assessment of learner outcome, product objectives; (3) a system for recording and reporting the results of process objective evaluation should be established in the school; and (4) product objectives and comparison group evaluation should take place at the end of the program. (DB)

ED 077973



PLANNING DOCUMENT IV

TM 002 783

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

PROCEDURES FOR PROGRAM EVALUATION

October, 1972

FILMED FROM BEST AVAILABLE COPY

ED 077973

TM 002 783

PROCEDURES FOR PROGRAM EVALUATION

Much of what has been said in the previous three planning documents has direct relationship to procedures for program evaluation. However, for the purpose of understanding the role of evaluation in planning it is necessary to understand only three types of evaluation.

The first type, called comparison group evaluation, is done apart from any performance objectives established for the program. Comparison evaluation simply means that two or more programs trying to accomplish similar ends are compared to see which one is most successful. In order to conduct this kind of evaluation, it is necessary to design the program implementation in such a way that there is random assignment of students to various competing programs.

It is interesting to note that comparison group evaluation can take place without the specification of performance objectives for a program. Unfortunately the performance objectives are too often thought of as merely an extension of evaluation, as opposed to their fundamental importance in program development.

The second type of program evaluation is the analysis of learner outcome or product objectives. Like the comparison group evaluation, the evaluation of learner outcome objectives reveals information on the overall success or failure of the program. However, the evaluation of these learner outcome objectives in isolation gives an absolute rather than a relative measure of success. For instance, two reading programs

that attained all of their learner outcome objectives might be ranked equally successful on this basis, but differ markedly if they could be compared one to another in a comparison group evaluation.

A third type of evaluation is related to the process objectives that are used to administer the program. Unlike the comparison evaluation and the product objectives evaluation, the process evaluation is related to the means of the program rather than the ends. To be of value, the process objectives are evaluated during the program's operation. Based on the outcome of process objective evaluation, projects can be modified in "midstream" in order to increase the probability of attaining outcome objectives.

The format for product and process objectives is to a degree a function of the type of measurement employed. Aside from objectives that are sometimes written in terms of a comparison group evaluation, most other kinds of performance objectives are written to fit one or more of the following evaluation formats:

	Examples (all product objectives)
1. An historical comparison - that is, objectives are written to compare a group's performance one year to the same group or a different group in some previous year.	The average score on the ABC test of reading comprehension will be higher for 1972 Grade 6 pupils than for 1971 Grade 6 pupils in Smith Elementary School.
2. Pre and post test comparisons - that is, objectives are written in terms of how much improvement is shown during a one year period by the same group.	The 6th grade pupils in Smith Elementary School will average one year gain in reading for the school year 1972-1973 as measured by the ABC test of reading.
3. Comparison to some norm - that is, objectives are written with standards of performance related to the results of norm referenced testing.	80% of the 6th grade students at Smith Elementary School will, by the end of the 1972-1973 school year, be reading at grade level as measured by the ABC test of reading.

4. Content based, mastery level comparison - that is, structuring the objectives so that the measurement is based on some prior judgments of mastery level. This kind of measurement is usually associated with criterion-referenced testing.

85% of the 6th grade students at Smith Elementary School will demonstrate survival literacy by scoring at the 35% correct level or better on a mastery test of literacy.

Hopefully the above discussion has convinced the reader that these three kinds of evaluation are all necessary to maximizing the information concerning program success or failure. Product objective and process objective evaluation can usually be conducted by program personnel. However, some assistance in specifying the performance objectives and creating record keeping documents for evaluation is usually needed. The comparison group evaluation, which need not be based upon the product or process objectives, is usually conducted by an evaluation staff. It is important that the evaluation staff is contacted early enough prior to program implementation to explore with the planning group the possibilities of implementing the program in such a way as to provide for comparison group evaluation.

The other important issue in evaluation is the type of measure or instrument employed. Measurements are either direct or indirect. In direct evaluation the behavior is completely observable. For example, the objective below can be evaluated in a very direct manner:

> All sixth grade students shall run the hundred yard dash in under twelve seconds.

In this example the physical skill involved in running a hundred yard dash under twelve seconds is directly observable and directly measured.

Most measurement in education usually is not of this direct type. Usually we gather a sample of behavior that represents some kind of internal activity. The typical achievement test in reading purports to indirectly sample a set of behaviors that are important to the total, internal act of reading.

Two kinds of indirect measures can be distinguished by whether or not the subject being measured actively or passively participates in the measurement act. In the reading test example above, the student is required to demonstrate his reading knowledge by actively taking the test. However, if we were interested in measuring something like reading interest, good interest tests might not be readily available. Therefore, it may be necessary to devise other kinds of indirect measures to pick up samples of behavior that might have implications for reading interest. For instance, voluntary trips to the library, a record of books checked out of the library, the amount of time spent in recreational reading, might all be observations from which one could infer a reading interest level. This latter kind of unobtrusive, indirect measure is the kind usually employed in trying to assess student self concept, reading interests, attitude towards school, and other aspects of affective behavior.

When it is determined that a test will be employed as a means of indirectly sampling some behavior, there are two basic kinds of tests employed. One is the standardized norm referenced test with which most educators are quite familiar. The Iowa Test of Basic Skills is a standardized, norm referenced test. The other type is called a criterion

referenced test, and while it quite often purports to measure the same thing as is measured by a norm referenced test, its construction, applicability, and resulting scores are quite different. There is not the time in this planning document to discuss the merits of each type of measurement except to suggest that the criterion referenced test is thought by some experts to be more applicable to planning. The reader is referred to the Department of Educational Research and Program Assessment for detailed information on norm referenced and criterion referenced testing.

This discussion of types of direct and indirect measurement is important to the person or group charged with the responsibility of developing performance objectives. Clearly, product, performance objectives need not be written solely based on some standardized test. The measurement of performance is limited only by the imagination of the objective writer and the limitations of measurement. The limitations of measurement are often ameliorated by combining measurement approaches. Each measurement approximates, and all approximations taken together usually provide more precision.

Higher order conceptual activity often is difficult to define in performance terms. However, this limitation applies to product objectives and not process objectives. Process objectives refer to how programs are implemented and are almost always amenable to description in performance terms. Even certain learner outcomes that initially are thought to be irreducable to performance description, can have some of their aspects

described in performance terms. For those learner outcomes that can't be measured, structured opinion-gathering instruments can be employed.

Once the evaluation type and instrumentation have been defined, there still is the question of what are the consequences of evaluation. Here, unfortunately, the specter of personal blame too often clouds the evaluation. It is not the purpose of program evaluation to ascribe fault to individuals or to reward others. While in some school systems, evaluation results have been used in this fashion, it is not the intent or desire of the evaluation component of the Milwaukee Public School planning process to be used in this way. The object of evaluation is to provide decision-making information about programs, not individuals. The consequence of not attaining desired student outcomes or not meeting process objectives, is to modify the program.

Steps in Program Evaluation:

1. Consideration of how objectives will be evaluated should begin prior to program development. Working through the administrative specialist, the principal and his planning staff may wish to consult Department of Educational Research and Program Assessment personnel on questions of evaluation. If the planning group desires comparison group evaluation, it is particularly important to contact evaluation personnel before program commitments have been made.

2. Following consultation with evaluation personnel, the planning group should make decisions on what measurement instruments, performance objective formats, and criterion levels represent fair and reasonable assessment of learner outcome, product objectives.

3. A system for recording and reporting the results of process objective evaluation should be established in the school. These evaluation results, received periodically throughout the year, can serve to modify the program while it is operating.

4. Product objectives and comparison group evaluation should take place at the end of the program. The results of these evaluations coupled with the implementation assessment gained from process objective evaluation, will form the basis for a new needs assessment.